Conformation changes and protein folding induced by ϕ^4 interaction

M. Januar

Department of Physics, University of Indonesia, Kampus UI Depok, Depok 16424, Indonesia

A. Sulaiman

Badan Pengkajian dan Penerapan Teknologi, BPPT Bld. II (19th floor), Jl. M.H. Thamrin 8, Jakarta 10340, Indonesia
E-mail: asulaiman@webmail.bppt.go.id, sulaiman@teori.fisika.lipi.go.id

L.T. Handoko

Group for Theoretical and Computational Physics, Research Center for Physics, Indonesian Institute of Sciences, Kompleks Puspiptek Serpong, Tangerang, Indonesia
E-mail: handoko@teori.fisika.lipi.go.id, handoko@fisika.ui.ac.id, laksana.tri.handoko@lipi.go.id

Department of Physics, University of Indonesia, Kampus UI Depok, Depok 16424, Indonesia

A model to describe the mechanism of conformational dynamics in protein based on matter interactions using lagrangian approach and imposing certain symmetry breaking is proposed. Both conformation changes of proteins and the injected non-linear sources are represented by the bosonic lagrangian with an additional ϕ^4 interaction for the sources. In the model the spring tension of protein representing the internal hydrogen bonds is realized as the interactions between individual amino acids and nonlinear sources. The folding pathway is determined by the strength of nonlinear sources that propagate through the protein backbone. It is also shown that the model reproduces the results in some previous works.

Keywords: protein dynamics, protein folding, lagrangian, ϕ^4 interaction

1. Introduction

The pathway of proteins are determined by the sequences of its amino acid constituents. The time ordered of protein folding sequence leads from the primary to the secondary and subsequent structures. The secondary structure consists of the shape representing each segment of a polypeptide tied by hydrogen bonds, van der Walls forces, electrostatic interaction and hydrophobic effects. It is moreover formed around a group of amino acids considered as the ground state. Then it is extended to include adjacent amino acids till the blocking amino acids are reached, and the whole protein chain along the polypeptide adopted its preferred secondary structure.

Our understanding on the underlying above-mentioned mechanism has unfortu-

nately not been at the satisfactory level. For instance, the studies based on statistical analysis of identifying the probabilities of locating amino acids in each secondary structure are still at the level of less than 75% accuracy. Moreover, the main mechanism responsible for a structured folding pathway have not yet been identified at all. On the other hand, it is known that the protein misfolding has been identified as the main cause of several diseases like cancers and so on.[1]

Recently, Mingaleev *et.al.* have shown that the nonlinear excitations play an important role in conformational dynamics by decreasing the effective bending rigidity of a biopolymer chain leading to a buckling instability of the chain.[2] Following this understanding, a model to explain the transition of a protein from a metastable to its ground conformation induced by solitons has been proposed.[3] In the model the mediator of protein transition is the Davydov solitons propagating through the protein backbone.

At present, the most reliable theoretical explanation for this kind of the conformational dynamics of biomolecules is the so-called ab initio quantum chemistry approach. This however requires astronomical computational power to deal with realistic biological systems.[4,5] In contrary, there are some phenomenological model describing the folding pathway as a result of the interplay between the energy transfer from a solitary solution that travels along the protein backbone and string tension.[6] There are also some attempts to describe the dynamics in term of elementary biomatter using field theory approach[7] and open quantum system.[8,9]

This paper follows the later approach, but starting from the first principle using the lagrangian method to derive the responsible interactions and to clarify its origins. The paper is organized as follows. First, the model and the underlying assumptions are explained in detail. It is then followed by the derivation of relevant equation of motions (EOMs). Summary and conclusion based on the numerical analysis are given at the end of the paper.

2. The models

The model is an extension of the toy model proposed in.[10] More than considering a self-interaction mechanism as proposed in[10] and subsequently developed in,[3,6] more realistic model is introduced. In the model, the dynamics of amino acids forming proteins is initially considered as a free and linear system of bosonic matters. Further, external nonlinear sources, like laser or light bunch, are introduced. The sources which propagate through the protein backbone interact each other with the amino acids to induce conformation changes.

The model describes the conformation changes as the dynamics of amino acids using a free and massive (relativistic) bosonic lagrangian as below,

$$l_c = \left(\partial_\mu \phi_c\right)^\dagger \left(\partial^\mu \phi_c\right) + \frac{1}{2} m_{\phi_c}^2 \phi_c^\dagger \phi_c \ , \tag{1}$$

where ϕ_c represents the conformation field. The hermite conjugate is $\phi^\dagger \equiv \left(\phi^*\right)^T$ for a general complex field ϕ. On the other hand, the nonlinear sources represented

by the field ϕ_s are also governed by a massless bosonic lagrangian,

$$l_s = (\partial_\mu \phi_s)^\dagger (\partial^\mu \phi_s) + V(\phi_s) \,, \tag{2}$$

with an additional potential $V(\phi_s)$ taking the typical ϕ^4- self-interaction,

$$V(\phi_s) = \frac{1}{4}\lambda (\phi_s^\dagger \phi_s)^2 \,, \tag{3}$$

where λ is the coupling constant. It should be noted that both scalar fields, $\phi_c = \phi_c(t,x)$ denotes the local curvature of the conformation at position x with $\phi_c(x) = 1$ or 0 for α or $\beta-$helix.

The choice of interactions in Eqs. (1) and (2) are justified by the following considerations,

- The conformation changes are assumed to be linear. It is actually not necessarily massive. Although one can put by hand the mass term $m^2_{\phi_c}\phi_c^\dagger\phi_c$ in the lagrangian as written above, the massive conformational field could also be generated dynamically through certain symmetry breaking as shown later.
- The source is assumed to be massless concerning the laser or light source injected to the protein chains to induce the foldings.
- Its non-linearity is realized by introducing the ϕ_s self-interaction which leads to the non-linear EOM.
- For the sake of simplicity, the lagrangian is imposed to be symmetry under certain transformations, for instance in the present case is time and parity symmetry, *i.e.* $\phi(t,x) \to -\phi(-t,-x)$ for one-dimensional space.

We should remark here that the model is although written in a relativistic form, after deriving relevant EOMs one can take its non-relativistic limits to obtain final EOMs describing the desired dynamics. Secondly, instead of using the vector electromagnetic field A_μ to represent the nonlinear sources, like laser for instance, it is more convenient to consider the nonlinear source as a bunch of light or laser such that one might represent it in a 'macrosocopic' scalar field ϕ_s.

Considering the dimensional counting and the invariance on time-parity symmetry, the most general interaction between the conformation field and nonlinear sources is,

$$l_{\rm int} = -\Lambda (\phi_c^\dagger \phi_c)(\phi_s^\dagger \phi_s) \,, \tag{4}$$

with Λ denotes the strength of the interaction. Eqs. (3) and (4) lead to the total potential in the model,

$$V_{\rm tot} = \frac{1}{4}\lambda (\phi_s^\dagger \phi_s)^2 - \Lambda (\phi_c^\dagger \phi_c)(\phi_s^\dagger \phi_s) \,. \tag{5}$$

Eqs. (1), (2) and (5) provide the underlying interactions in the model.

Concerning the minima of the total potential in term of source field, that is

$$\left.\frac{\partial V_{\rm tot}}{\partial \phi_s}\right|_{\langle\phi_s\rangle,\langle\phi_c\rangle} = 0 \,, \tag{6}$$

at the vacuum expectation values (VEV) of the fields yields the non-trivial solution,

$$\langle \phi_s \rangle = \sqrt{\frac{2\Lambda}{\lambda}} \langle \phi_c \rangle \ . \tag{7}$$

Imposing certain local symmetry, namely the phase or U(1) symmetry to the above total lagrangian, the VEV in Eq. (7) obviously breaks the symmetry. The symmetry breaking at the same time shifts the mass term for ϕ_c as follow,

$$m_{\phi_c}^2 \to \overline{m}_{\phi_c}^2 \equiv m_{\phi_c}^2 - \frac{2\Lambda^2}{\lambda} \langle \phi_c \rangle^2 \ , \tag{8}$$

from Eq. (4).

On the other hand, Eq. (7) induces the 'tension force' which plays an important role to enable folded pathway in the present model. This will be discussed in the following section.

3. EOMs and its behaviours

Having the total lagrangian at hand, one can derive the EOM's using the Euler-Lagrange equation,

$$\frac{\partial l_{\rm tot}}{\partial \phi} - \partial_\mu \frac{\partial l_{\rm tot}}{\partial (\partial_\mu \phi)} = 0 \ , \tag{9}$$

where $l_{\rm tot} = l_c + l_s + l_{\rm int}$.

Substituting Eqs. (1), (2) and (4) into Eq. (9) in term of ϕ_c and ϕ_s, one immediately obtains a set of EOMs,

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{1}{\hbar^2} m_{\phi_c}^2 c^2 + 2\Lambda \, \phi_s^2 \right) \phi_c = 0 \ , \tag{10}$$

$$\left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} + 2\Lambda \, \phi_c^2 - 3\lambda \, \phi_s^2 \right) \phi_s = 0 \ . \tag{11}$$

Here the natural unit is restored to make the light velocity c and $\hbar$ reappear in the equation.

The last term in Eq. (11) determines the non-linearity of the EOM of source. One should also put an attention in the last term of Eq. (10), *i.e.* $\sim k \, \phi_c$ with $k \sim 2\Lambda \langle \phi_s \rangle^2$. This actually induces the tension force in the dynamics of conformational field enabling the folded pathway as expected.

Hence, solving both EOMs in Eqs. (10) and (11) simultaneously would provide the contour of conformational changes in term of time and one-dimensional space components.

4. Numerical analysis

Since the EOMs under consideration involves non-linear term, one should solve them numerically. The numerical analysis and simulation in the present paper are done using the finite difference method.[11] Throughout numerical works, non-relativistic

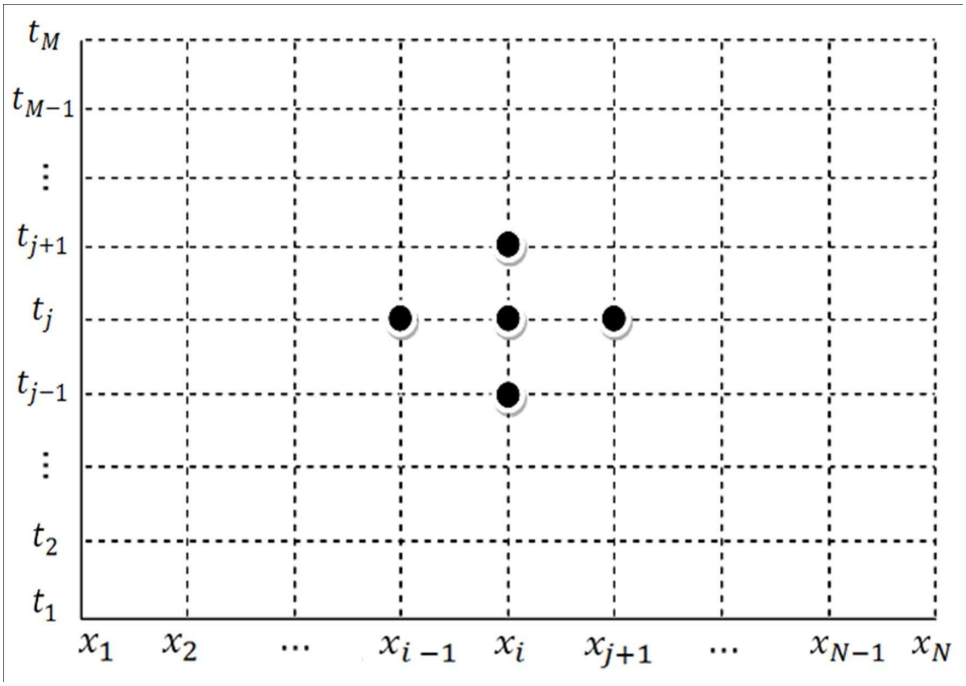

Fig. 1. The discretized grid for solving the EOMs over the coordinate space R.

limit $v = \partial x/\partial t \ll c$ and the following boundary conditions for both fields are deployed,

$$\begin{aligned} &\phi_s(0,t) = \phi_s(L,t) = 0 \ \text{ and } \ \phi_c(0,t) = \phi_c(L,t) = 0 && \text{for } 0 \le t \le b\ , \\ &\phi_s(x,0) = f(x) \ \text{ and } \ \phi_c(x,0) = p(x) && \text{for } 0 \le x \le L\ , \\ &\frac{\partial \phi_s(x,0)}{\partial t} = g(x) \ \text{ and } \ \frac{\partial \phi_c(x,0)}{\partial t} = q(x) && \text{for } 0 < x < L\ , \end{aligned} \tag{12}$$

with $f(x)$, $p(x)$, $g(x)$ and $q(x)$ are newly introduced auxiliary functions. In finite difference scheme, it is more convenient to replace ϕ_s and ϕ_c with u and w respectively, and rewrite them in discrete forms. Then, let us consider the coordinate space $R = \{(x,t) : 0 \le x \le L, 0 \le t \le b\}$ discretized on a grid consisting of $(N-1) \times (M-1)$ rectangles with side length $\Delta x = \delta$ and $\Delta t = \epsilon$ shown in Fig. 1. Solving the equations over the grid gives us the desired numerical solutions.

Both coupled EOMs in Eqs. (10) and (11) are rewritten in explicit discrete forms

as follows,

$$u_{i,j+1} = 2u_{i,i} - u_{i,j-1} + c^2\epsilon^2\left(\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{\delta^2} + 2\Lambda w_{i,j}^2 u_{i,j} - 3\lambda u_{i,j}^3\right) \quad (13)$$

$$w_{i,j+1} = 2w_{i,i} - w_{i,j-1} + c^2\epsilon^2\left(\frac{w_{i+1,j} - 2w_{i,j} + w_{i-1,j}}{\delta^2} + 2\Lambda u_{i,j}^2 w_{i,j} - \frac{c^2}{\hbar^2}m_{\phi_c}^2 w_{i,j}\right), \quad (14)$$

for $i = 2, 3, \cdots, N-1$ and $j = 2, 3, \cdots, M-1$. In order to calculate all values of Eqs. (13) and (14), the initial values for two lowest rows in Fig. 1 must be given. On the other hand, the value at t_1 is fixed by the boundary conditions in Eq. (12). The second order of Taylor expansion can also be used to determine the values in the second row. Therefore, the values at t_2 are determined by,

$$u_{i,2} = f_i - \epsilon g_i + \frac{c^2\epsilon^2}{2}\left(\frac{f_{i+1} - 2f_i + f_{i-1}}{\delta^2} + 2\Lambda p_i^2 f_i - 3\lambda f_i^3\right), \quad (15)$$

$$w_{i,2} = p_i - \epsilon q_i + \frac{c^2\epsilon^2}{2}\left(\frac{p_{i+1} - 2p_i + p_{i-1}}{\delta^2} + 2\Lambda f_i^2 p_i - \frac{c^2}{\hbar^2}m_{\phi_c}^2 p_i\right), \quad (16)$$

for $i = 2, 3, \cdots, N-1$.

For the initial stage, suppose the nonlinear sources has a particular form $f(x) = 2\mathrm{sech}(2x)\,\mathrm{e}^{i2x}$ and $g(x) = 1$ to generate the α-helix, while $g(x) = q(x) = 0$ for the sake of simplicity. Then, one can obtain the initial values in this case using Eqs. (15) and (16). The subsequent values are generated by substituting the preceeding values into Eqs. (13) and (14). The higher order values can be obtained using iterative procedure.

The result is given in Fig. 2. The left figure in each box describes the propagation of nonlinear sources in protein backbone, while the right one shows how the protein is folded. As can be seen in the figure, the protein backbone is initially linear before the nonlinear source injection. As the soliton started propagating over the backbone, the conformational changes appear. It should be remarked that the result is obtained up to the second order accuracy in Taylor expansion. In order to guarantee that the numerical solutions do not contain large amount of truncation errors, the step sizes δ and ϵ are kept small enough. Nevertheless, this should be good approximation to describe visually the mechanism of protein folding.

5. Conclusion

An extension of phenomenological model describing the conformational dynamics of proteins is proposed. The model based on the matter interactions among the relevant constituents, namely the conformational field and the nonlinear sources represented as the bosonic fields ϕ_c and ϕ_s. It has been shown that from the relativistic bosonic lagrangian with ϕ_s^4 self-interaction, the nonlinear and tension force terms appear naturally as expected in some previous works.[6]

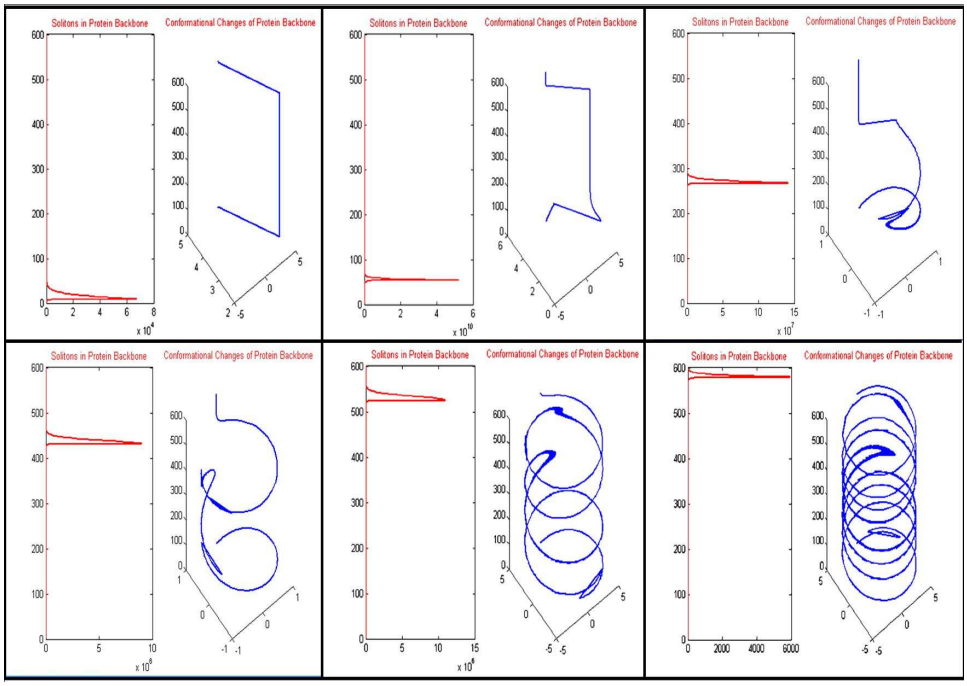

Fig. 2. The soliton propagations and conformational changes on the protein backbone inducing protein folding. The vertical axis in soliton evolution denotes time in second, while the horizontal axis denotes its amplitude. The conformational changes are on the (x, y, z) plane.

However, the present model has different contour since the EOMs governing the whole dynamics are the linear and nonlinear Klein-Gordon equations. Note that the original model by Berloff deployed the linear Klein-Gordon and nonlinear Schrodinger equations.

Moreover, the present model has inhomogenous tension force, in contrast with the homogeneous tension force in the Berloff's model, due to simultaneous solutions of Eqs. (10) and (11). These lead to wrigling folded pathways as shown in Fig. 2 which should be more natural than the homogeneous one.

Acknowledgments

The authors greatly appreciate fruitful discussion with T.P. Djun throughout the work. AS thanks the Group for Theoretical and Computational Physics LIPI for warm hospitality during the work. This work is partially funded by the Indonesia Ministry of Research and Technology and the Riset Kompetitif LIPI in fiscal year 2010 under Contract no. 11.04/SK/KPPI/II/2010.

References

1. C. M. Dobson, *Nature* **426**, p. 884 (2003).
2. S. F. Mingaleev, Y. B. Gaididei, P. L. Christiansen and Y. Kivshar, *Europhys. Lett.* **59**, p. 403 (2002).
3. S. Caspi and E. Ben-Jacob, *Phys. Lett. A* **272**, p. 124 (2000).
4. A. Garcia and J. Onuchi, *Proc. Natl. Acad. Sci. USA* **100**, p. 13898 (2003).
5. J. N. Onuchi and P. G. Wolynes, *Curr. Opin. Struct. Biol.* **14**, p. 70 (2004).
6. N. G. Berloff, *Phys. Lett. A* **337**, p. 391 (2005).
7. A. Sulaiman and L. T. Handoko, *J. Compt. Theor. Nanoscience* **in press** (2010).
8. A. Sulaiman, F. P. Zen, H. Alatas and L. T. Handoko, *Phys. Rev. E* **81**, p. 061907 (2010).
9. A. Sulaiman, F. P. Zen, H. Alatas and L. T. Handoko, *Int. J. Mod. Phys. A* **in press** (2010).
10. S. Caspi and E. Ben-Jacob, *Europhys. Lett.* **47**, p. 522 (1999).
11. J. H. Mathews and K. D. Fink, *Numerical Methods using Matlab 4th Ed.* (Prentice-Hall, 2004).